

#11

1638

## RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/578,827A

TIME: 15:15:08

Input Set : A:\5914-078-999.txt

Output Set: N:\CRF3\09102001\I578827A.raw

ENTERED

```

5 <110> APPLICANT: Benfey, Philip
7     Helariutta, Yrjo
9     Fukaki, Hidehiro
11    Nakajima, Keiji
15 <120> TITLE OF INVENTION: SHORT-ROOT GENE, PROMOTER, AND USES THEREOF
19 <130> FILE REFERENCE: 5914-078-999
23 <140> CURRENT APPLICATION NUMBER: 09/578,827A
25 <141> CURRENT FILING DATE: 2000-05-24
29 <160> NUMBER OF SEQ ID NOS: 17
33 <170> SOFTWARE: PatentIn version 3.0
37 <210> SEQ ID NO: 1
39 <211> LENGTH: 2825
41 <212> TYPE: DNA
43 <213> ORGANISM: Arabidopsis thaliana
47 <400> SEQUENCE: 1
48 atcgattaag agaaaataga gttttcatgc accagtgttg atagtaacgt agtcgcggaa      60
50 tgtctaaaac gattatgagt ttggtgtttt gattggtttag aattgggtatt agtaggacat      120
52 tctaactttt ttgtagtctt gttgatttag gatgcgtaaa gagtcttttt attttacacc      180
54 agttgagact tgggatcgat agtacttgaa acacttggtt ggtttcatgt atttggccta      240
56 tatataaaca aacatcgtaa ttatatacgg atttttttcg gaattttacg ccatatctgt      300
58 aagtatatat aacatgcatt tcgttttcaa attcatatga tgaacgatcc acgtaagtgc      360
60 tactactcct acaatattgc atgagagaga tatgtattta taaattttat tttgaagaag      420
62 aaataagagg gaaggttact tgggtggatc gatgtgaaaa caaaagaaga aaaagcgaaa      480
64 cccactaagc cattacatga tatcgacctt cttatctttt tcctctttat tttatttttc      540
66 tcaggacttt tttctactta atgaaacctc caaactatct aactaatata ctcccatgta      600
68 gaataaagaa aattatataa gatattgttg atattttgta actagaaaaa atatttgctc      660
70 tgtaattttt cgtaagttaa atcaacattt ttcagtagaa acaaattatta ctgcaaaaag      720
72 taggatcatt atttttgtcc aaaatctcag ttagctatag ggtttagta aaaacaaaac      780
74 acattcttga tttgccccaa aaaataaaga gagagaagaa tattgttcaa aagtgggtctc      840
76 ttctctctct aattatgttt tcaactaaacc caattagatt caaacagtct acaaagtcca      900
78 aaagataaac atgggacaac aattcgatgc aaaaaatcct cttttcatgc tcttttttta      960
80 ttctctagtc ttttaaatta ctaataaaaa ctcacaaatc caccaaacc cttctctaca      1020
82 actcaccttc atctagattt acccactccc accgagaaac acaagaaaaa aaatatacat      1080
84 atataaatat acaagacaac acatgatgct gatgcaatat acacaacaaa gtattaaatc      1140
86 ttagatattg tgggtctccc tttcttctat tcattttctt attcattaaa aaaaaaaaaa      1200
88 ggatactctc ttttagactag tcagtctcca acaacaacaa caatccgata gtatcattac      1260
90 aaatcaatct tcgttaagca gaacttccac caccactact ggctctccac aaactgctta      1320
92 tcaactaac tttccacaaa acgacgtcgt cgaagaatgc ttcaactttt tcatggatga      1380
94 agaagacctt tcctcttctt cttctcacca caaccatcac aaccacaaca atcctaatac      1440
96 ttactactct cttttcacta ctcccaccca ataccatccc gccacatcat caacccttc      1500
98 ctccaccgcc gcagccgcag ctttagcctc gccttactcc tcctccggcc accataatga      1560
100 ccttccgcgc tttccatac ctcaaacctc tcgctccttc gacttctcag ccaatgccaa      1620
102 gtgggcagac tcggtccttc ttgaagcggc acgtgccttc tccgacaaag acactgcacg      1680
104 tgcgcaacaa atcctatgga cgtcacaaga gctctcttct ccgtacggag acaccgagca      1740
106 aaaactggct tcttacttcc tccaagctct cttcaaccgc atgaccgggt caggcgaaacg      1800
108 atgtaccga accatggtaa cagctgcagc cacagagaag acttgctcct tcgagtcaac      1860
110 gcgaaaaact gtactaaagt tccaagaagt tagcccctgg gccacgtttg gacacgtggc      1920

```

## RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/578,827A

TIME: 15:15:08

Input Set : A:\5914-078-999.txt

Output Set: N:\CRF3\09102001\I578827A.raw

```

112 ggcaaacgga gcaatcttgg aagcagtaga cggagaggca aagatccaca tcgttgacat 1980
114 aagctccacg ttttgcactc aatggccgac tcttctagaa gcttttagcca caagatcaga 2040
116 cgacacgcct cacctaaggc taaccacagt tgtcgtggcc aacaagtttg tcaacgatca 2100
118 aacggcgctc catcggatga tgaaagagat cggaaaccga atggagaaat tcgctaggct 2160
120 tatgggagtt cttttcaaat ttaacattat tcatcacgtt ggagatttat ctgagtttga 2220
122 tctcaacgaa ctcgacgtta aaccagacga agtccttggcc attaactgcg taggcgcgat 2280
124 gcatgggagc gcttcacgtg gaagccctag agacgctgtg atatcgagtt tccgacggtt 2340
126 aagaccgagg attgtgacgg tcgtagaaga agaagctgat cttgtcggag aagaagaagg 2400
128 tggctttgat gatgagttct tgagagggtt tggagaatgt ttacgatggt ttagggtttg 2460
130 cttcgagtca tgggaagaga gttttccaag gacgagcaac gagaggttga tgctagagcg 2520
132 tgcagcggga cgtgcgatcg ttgatcttgt ggcttgtgag ccgtcggatt ccacggagag 2580
134 gcgagagaca gcgagggaagt ggtcgaggag gatgaggaat agtgggtttg gagcgggtggg 2640
136 gtatagtgat gaggtggcgg atgatgtcag agctttgttg aggagatata aagaaggtgt 2700
138 ttggtcgatg gtacagtgtc ctgatgccgc cggaatatcc ctttgtttga gagatcagcc 2760
140 ggtggttttg gctagtgcgt ggcggccaac gtaaagggtt gtttttattt tttcataagg 2820
142 aattc 2825
145 <210> SEQ ID NO: 2
147 <211> LENGTH: 531
149 <212> TYPE: PRT
151 <213> ORGANISM: Arabidopsis thaliana
155 <400> SEQUENCE: 2
157 Met Asp Thr Leu Phe Arg Leu Val Ser Leu Gln Gln Gln Gln Ser
158 1 5 10 15
160 Asp Ser Ile Ile Thr Asn Gln Ser Ser Leu Ser Arg Thr Ser Thr Thr
161 20 25 30
163 Thr Thr Gly Ser Pro Gln Thr Ala Tyr His Tyr Asn Phe Pro Gln Asn
164 35 40 45
166 Asp Val Val Glu Glu Cys Phe Asn Phe Phe Met Asp Glu Glu Asp Leu
167 50 55 60
169 Ser Ser Ser Ser Ser His His Asn His His Asn His Asn Asn Pro Asn
170 65 70 75 80
172 Thr Tyr Tyr Ser Pro Phe Thr Thr Pro Thr Gln Tyr His Pro Ala Thr
173 85 90 95
175 Ser Ser Thr Pro Ser Ser Thr Ala Ala Ala Ala Leu Ala Ser Pro
176 100 105 110
178 Tyr Ser Ser Ser Gly His His Asn Asp Pro Ser Ala Phe Ser Ile Pro
179 115 120 125
181 Gln Thr Pro Pro Ser Phe Asp Phe Ser Ala Asn Ala Lys Trp Ala Asp
182 130 135 140
184 Ser Val Leu Leu Glu Ala Ala Arg Ala Phe Ser Asp Lys Asp Thr Ala
185 145 150 155 160
187 Arg Ala Gln Gln Ile Leu Trp Thr Leu Asn Glu Leu Ser Ser Pro Tyr
188 165 170 175
190 Gly Asp Thr Glu Gln Lys Leu Ala Ser Tyr Phe Leu Gln Ala Leu Phe
191 180 185 190
193 Asn Arg Met Thr Gly Ser Gly Glu Arg Cys Tyr Arg Thr Met Val Thr
194 195 200 205
196 Ala Ala Ala Thr Glu Lys Thr Cys Ser Phe Glu Ser Thr Arg Lys Thr
197 210 215 220

```

## RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/578,827A

TIME: 15:15:08

Input Set : A:\5914-078-999.txt

Output Set: N:\CRF3\09102001\I578827A.raw

```

199 Val Leu Lys Phe Gln Glu Val Ser Pro Trp Ala Thr Phe Gly His Val
200 225                230                235                240
202 Ala Ala Asn Gly Ala Ile Leu Glu Ala Val Asp Gly Glu Ala Lys Ile
203                245                250                255
205 His Ile Val Asp Ile Ser Ser Thr Phe Cys Thr Gln Trp Pro Thr Leu
206                260                265                270
208 Leu Glu Ala Leu Ala Thr Arg Ser Asp Asp Thr Pro His Leu Arg Leu
209                275                280                285
211 Thr Thr Val Val Val Ala Asn Lys Phe Val Asn Asp Gln Thr Ala Ser
212    290                295                300
214 His Arg Met Met Lys Glu Ile Gly Asn Arg Met Glu Lys Phe Ala Arg
215 305                310                315                320
217 Leu Met Gly Val Pro Phe Lys Phe Asn Ile Ile His His Val Gly Asp
218                325                330                335
220 Leu Ser Glu Phe Asp Leu Asn Glu Leu Asp Val Lys Pro Asp Glu Val
221                340                345                350
223 Leu Ala Ile Asn Cys Val Gly Ala Met His Gly Ile Ala Ser Arg Gly
224                355                360                365
226 Ser Pro Arg Asp Ala Val Ile Ser Ser Phe Arg Arg Leu Arg Pro Arg
227    370                375                380
229 Ile Val Thr Val Val Glu Glu Glu Ala Asp Leu Val Gly Glu Glu Glu
230 385                390                395                400
232 Gly Gly Phe Asp Asp Glu Phe Leu Arg Gly Phe Gly Glu Cys Leu Arg
233                405                410                415
235 Trp Phe Arg Val Cys Phe Glu Ser Trp Glu Glu Ser Phe Pro Arg Thr
236                420                425                430
238 Ser Asn Glu Arg Leu Met Leu Glu Arg Ala Ala Gly Arg Ala Ile Val
239                435                440                445
241 Asp Leu Val Ala Cys Glu Pro Ser Asp Ser Thr Glu Arg Arg Glu Thr
242    450                455                460
244 Ala Arg Lys Trp Ser Arg Arg Met Arg Asn Ser Gly Phe Gly Ala Val
245 465                470                475                480
247 Gly Tyr Ser Asp Glu Val Ala Asp Asp Val Arg Ala Leu Leu Arg Arg
248                485                490                495
250 Tyr Lys Glu Gly Val Trp Ser Met Val Gln Cys Pro Asp Ala Ala Gly
251                500                505                510
253 Ile Phe Leu Cys Trp Arg Asp Gln Pro Val Val Trp Ala Ser Ala Trp
254    515                520                525
256 Arg Pro Thr
257    530

```

259 &lt;210&gt; SEQ ID NO: 3

261 &lt;211&gt; LENGTH: 2059

263 &lt;212&gt; TYPE: DNA

265 &lt;213&gt; ORGANISM: Arabidopsis thaliana

269 &lt;400&gt; SEQUENCE: 3

```

270 aaaaaaaaaa aatggatact ctcttttagac tagtcagtct ccaacaacaa caacaatccg      60
272 atagtatcat tacaaatcaa tcttcgttaa gcagaacttc caccaccact actggctctc      120
274 cacaaactgc ttatcactac aactttccac aaaacgacgt cgtcgaagaa tgcttcaact      180
276 ttttcatgga tgaagaagac ctttctctct cttcttctca ccacaaccat cacaaccaca      240

```

## RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/578,827A

TIME: 15:15:08

Input Set : A:\5914-078-999.txt

Output Set: N:\CRF3\09102001\I578827A.raw

```

278 acaatcctaa tacttactac tctcctttca ctactccac ccaataccat cccgccacat 300
280 catcaacccc ttctccacc gccgcagccg cagcttttagc ctgcgcttac tctcctccg 360
282 gccaccataa tgacccttcc gcggttctcca tacctcaaac tctcgcgtcc ttgcacttct 420
284 cagccaatgc caagtgggca gactcggtcc ttcttgaagc ggcacgtgcc ttctccgaca 480
286 aagacactgc acgtgcgcaa caaatcctat ggacgctcaa cgagctctct tctccgtaat 540
288 gaaaaccgct tcattttcct tgtatttgtc tgaggttagg attagaccat tggttgttac 600
290 tttcgaattc ttccaattta gttgttactt tcgaattctt ccatctctta' gtttactaaa 660
292 acaaacttat gtgccccata tttctccaac aatttgttga gtggtagctt acgttttact 720
294 gtatacgctt ttgcaggtta tatcagcaca accattaatg atggcccggg atgtttgatg 780
296 ctaagatgtc ctgaccatc ttgtcttgct gctgttggtc atgatatggt tgacaaatta 840
298 gcgctctgaag acgaaaagga gaagtacaac agatattttc ttaggtctta tattgaagac 900
300 aacagaaagg taagcagttc agaaaattta tatcacacag actggtatta atgtcgctgg 960
302 tcttttattg agcaaaaact ggcttcttac ttcttccaag ctctcttcaa ccgcattgacc 1020
304 ggttcaggcg aacgatgcta ccgaaccatg gtaacagctg cagccacaga gaagacttgc 1080
306 tccttcgagt caacgcgaaa aactgtacta aagttccaag aagttagccc ctgggccacg 1140
308 tttggacacg tggcgccaaa cggagcaatc ttggaagcag tagacggaga ggcaaagatc 1200
310 cacatcgttg acataagctc caggttttgc actcaatggc cgactcttct agaagcttta 1260
312 gccacaagat cagacgacac gccacaccta aggtcaacca cagttgtcgt ggccaacaag 1320
314 tttgtcaacg atcaaacggc gtgcacatcg atgatgaaag agatcggaaa ccgaatggag 1380
316 aaattcgcta ggcttatggg agttccttcc aaatttaaca ttattcatca cgttggagat 1440
318 ttatctgagt ttgatctcaa cgaactcgac gttaaaccag acgaagtctt ggccattaac 1500
320 tgcgtaggcg cgatgcattg gatcgcttca cgtggaagcc ctagagacgc tgtgatatcg 1560
322 agtttccgac ggttaagacc gaggattgtg acggtcgtag aagaagaagc tgatcttgct 1620
324 ggagaagaag aaggtggctt tgatgatgag ttcttgagag ggtttggaga atgtttacga 1680
326 tggtttaggg tttgcttcga gtcattggaa gagagtttcc caaggacgag caacgagagg 1740
328 ttgatgctag agcgtgcagc gggacgtgcg atcgttgatc ttgtggcttg tgagccgtcg 1800
330 gattccacgg agaggcgaga gacagcgagg aagtggctga ggaggatgag gaatagtggg 1860
332 tttggagcgg tgggggtatg tgatgaggtg gcggatgatg tcagagcttt gttgaggaga 1920
334 tataaagaag gtgtttggtc gatggtacag tgcctgatg ccgccggaat attcctttgt 1980
336 tggagagatc agccggtggt ttgggctagt gcgtggcgcc caacgtaaaag ggttgttttt 2040
338 attttttcat aaggaattc 2059
341 <210> SEQ ID NO: 4
343 <211> LENGTH: 2502
345 <212> TYPE: DNA
347 <213> ORGANISM: Arabidopsis thaliana
351 <400> SEQUENCE: 4
352 agaagcagag cgtggggttt cttctaataa ttgtagaaga aactgatcat gagaacattt 60
354 gatctaccag agatggtgat gactcataag atgtaaatat ctactgcatt atgtctagcc 120
356 taggtataaa tgtagatttg atcactttct tcattaatta gtttggaatt ttagcatgat 180
358 atagcatata tctaaatatg tccgaaactt tctacatac tagaaaatat ggagagttat 240
360 gtaatgtagg tttgcttggt aatatacaaa ataacatcat catttagttt ttagattttt 300
362 tattttattt tttataatgg tgctacgtac gtggcgatca aattattcca attttgagac 360
364 ttcgggattt taaacgaaat taaacaatgg gcatgagctc ggggggatag acaagattaa 420
366 tgctttgtat cgagacaaac gagaaaatca tgatgagcct atgcattaag tgccgttggt 480
368 taattagagg ttcgcatata cataaaccag tagacatatg gataaatatg aacacacaca 540
370 ccaaaaaagt gggaaatcta aataagtgtg gagaataata agtcctcagg tgggagattc 600
372 aaagagagga caatgaaggg tatatagact ctaaacaaaa atggcatgac ttagtggaga 660
374 gggtttttaa ttgaaacaag taggattgaa gaacaagaaa acaaagaagc atgccctaga 720
376 tttctgagat aataattaca cattgctgtt tatataaggt aagagaatat gacacattgg 780

```

## RAW SEQUENCE LISTING

DATE: 09/10/2001

PATENT APPLICATION: US/09/578,827A

TIME: 15:15:08

Input Set : A:\5914-078-999.txt

Output Set: N:\CRF3\09102001\I578827A.raw

```

378 ttggttttctt acgggtaaat gtgaagaaaa aaaaatagta atatttgaga aaatctaaaa 840
380 tagtaaagag gtatatatgg agaagaagag agaaaagggg aaaaatagtg cagagaatgg 900
382 agagagggtta ggaggcaaag gcaaatgtgg agctttgatg atggtgatgc acgccgtcag 960
384 cttttcttca cgctgtctcc cactcactca cacctatgaa cattctctct ctattttata 1020
386 attatattca catgtctcta tgttactatg taaatggtga ccacttaagt atttatatat 1080
388 catgtatata tcttataggt atcatacaaa atggtcatga aacttttgca atttcaatct 1140
390 acttggtcat tgtagatgct agcttttcac atgttttgaa aattagtcct gatctgaaat 1200
392 tctttaatta gcattgtttt gttggtcaac gtttaatttc ttgattattg atgtcaaaaa 1260
394 ttcagagcgt tcagaactct tacactaatt tcttaaaaaa aatcgattaa gagaaaatag 1320
396 agttttcatg caccagtgtt gatagtaacg tagtcgcgga atgtctaaaa cgattatgag 1380
398 tttggtgttt tgattggtta gaattggtat tagtaggaca ttctaacttt tttgtagtct 1440
400 tgttgattta ggatgcgtaa agagtctttt tattttacac cagttgagac ttgggatcga 1500
402 tagtacttga aacacttggt tggtttcatg tattttggct atatataaac aaacatcgta 1560
404 attatatacg gatttttttc ggaattttac gccatatctg taagtatata taacatgcat 1620
406 gtcgttttca aattcatatg atgaacgac cagtaagtgc ctactactcc tacaatattg 1680
408 catgagagag atatgtattt ataaatttta ttttgaagaa gaaataagag ggaaggttac 1740
410 ttgggtggat cgatgtgaaa acaaaagaag aaaaagcgaa acccactaag ccattacatg 1800
412 atatcgacct tcttatcttt ttctctttta ttttattttt ctcaggactt ttttctactt 1860
414 aatgaaacct ccaaactatc taactaatac actcccatgt agaataaaga aaattatata 1920
416 agatattggt gatattttgt aactagaaaa tatatttgct ctgtaatttt tcgtaagtta 1980
418 aatcaacatt tttcagtaga aacaaatatt actgcaaaaa gtaggatcat tatttttgct 2040
420 caaaatctca gttagctata gggttgtagt aaaaacaaaa cacattcttg atttgcccca 2100
422 aaaaataaag agagagaaga atattgttca aaagtgtctt cttctctctc taattatggt 2160
424 ttcactaaac ccaattagat tcaaacagtc tacaagtc aaagataaa catgggacaa 2220
426 caattogatg caaaaaatcc tcttttcatg ctcttttttt attctctagt cttttaaatt 2280
428 actaataaaa actcacaat ccaccaaacc cattctctac aactcacctt catctagatt 2340
430 taccactcc caccgagaaa cacaagaaaa aaaatataca tatataaata tacaagacaa 2400
432 cacatgatgc tgatgcaata tacacaacaa agtattaaat cttagatatt gtgggtctcc 2460
434 ctttcttcta ttcattttct tattcattaa aaaaaaaaaa tg 2502

```

437 &lt;210&gt; SEQ ID NO: 5

439 &lt;211&gt; LENGTH: 58

441 &lt;212&gt; TYPE: PRT

443 &lt;213&gt; ORGANISM: Arabidopsis thaliana

447 &lt;400&gt; SEQUENCE: 5

449 Pro Phe Ile Arg Phe Thr Gln Leu Thr Ala Asn Gln Ala Ile Leu Glu

450 1 5 10 15

452 Ala Ile Asn Gly Asn His Gln Ala Ile His Ile Val Asp Phe Asp Ile

453 20 25 30

455 Asn His Gly Val Gln Trp Pro Pro Leu Met Gln Ala Leu Ala Asp Arg

456 35 40 45

458 Tyr Pro Ala Pro Thr Leu Arg Ile Thr Gly

459 50 55

461 &lt;210&gt; SEQ ID NO: 6

463 &lt;211&gt; LENGTH: 58

465 &lt;212&gt; TYPE: PRT

467 &lt;213&gt; ORGANISM: Arabidopsis thaliana

471 &lt;400&gt; SEQUENCE: 6

473 Pro Tyr Leu Lys Phe Ala His Phe Thr Ala Asn Gln Ala Ile Leu Glu

474 1 5 10 15

**VERIFICATION SUMMARY**

**PATENT APPLICATION: US/09/578,827A**

**DATE: 09/10/2001**

**TIME: 15:15:09**

**Input Set : A:\5914-078-999.txt**

**Output Set: N:\CRF3\09102001\I578827A.raw**